

Modeling Typological Markedness in Semantics: The Case of
Negative Concord

Frank Richter and Manfred Sailer

University of Tübingen and University of Göttingen

Proceedings of the HPSG06 Conference

Linguistic Modelling Laboratory
Institute for Parallel Processing
Bulgarian Academy of Sciences
Sofia
Held in Varna

Stefan Müller (Editor)

2006

CSLI Publications

<http://csli-publications.stanford.edu/>

Abstract

We will provide an analysis of negative concord in sentential negation in three languages, French, Polish and German. The focus of the paper is (1) the typological variation with respect to the realization of negative concord in the three languages under investigation and (2) the treatment of lexical exceptions within the different typological classes. We will propose a unified theory of negative concord which identifies a common core system and adds language-specific constraints which can handle typological variation between languages and lexical exceptions within a given language.

1 Introduction

Negative concord (NC) can be explored from two perspectives: Either the general pattern of negation is investigated from a typological perspective or the negation system of a particular language is presented in considerable detail. In this study we attempt to combine these two approaches and propose a fine-grained analysis including idiosyncratic exceptions embedded in a typological perspective. A key insight for our analysis is the observation that languages which are predominantly NC languages often contain lexical exceptions to this tendency, i.e. words which do not enter into a concord relationship. Similarly, languages in which multiple negative expressions are obligatorily interpreted as separate negative quantifiers ($\neg\exists$) may contain words which prefer a negative concord interpretation. We consider it an important feature of our theory that it is formulated in a surface-oriented framework without abstract syntactic nodes or invisible categories which drive the semantic interpretation.

We will present a grammar architecture for expressing the difference between optional, obligatory and impossible NC as a consequence of different realizations of cross-linguistic properties of language. Our typological approach, illustrated here with data from French, Polish and German, aims at modeling NC across languages as a consequence of different basic principles of the semantic combinatorics, of language-specific constraints, and of idiosyncratic lexical properties. This lexicalist view will be supported with lexical items in languages with predominantly obligatory concord or predominantly impossible concord which break the general pattern and can only be described as lexicalized exceptions. This will lead to a theory which is prepared to accommodate exceptions without imposing mutually inconsistent constraints. We will argue that the basic principles should be expressed in terms of agreement requirements and the lexical idiosyncrasies as collocational restrictions.

[†]We thank Olivier Bonami, Gilles Boyé, and Danièle Godard for comments and creative suggestions on French, Beata Trawiński and Adam Przepiórkowski for help with Polish, and Garrett Hubing for proofreading the manuscript.

2 French, Polish and German

In this section we give a brief overview of the distribution of n-words and sentential negation markers as well as their interaction in producing sentential negation in finite sentences in French, Polish and German. Although sentential negation in French and German will (in accordance with the literature) be identified as belonging to different typological systems, we will then proceed to show that there are exceptional expressions in both languages with apparently identical syntactic distribution and semantic behavior.

2.1 Core Data

Negation in French, which is a standard example of an NC language, has been studied thoroughly (Gaatone, 1971; Cristea, 1971; Muller, 1991; Grevisse and Goosse, 1993). It is famous for the peculiar behavior of the lexical elements that are associated with negation. The most prominent ones are the pre-verbal negation particle *ne*, the negative adverb *pas*, and so-called n-words such as *personne* (*nobody*) and *rien* (*nothing*). We will follow the syntactic analysis of French negation as proposed in Kim (1996). As in Rowlett (1998) we suppose that the pre-verbal negation particle *ne* does not carry semantic negation in any register of modern European French, but the negative adverb *pas* always does.¹ On the basis of these two assumptions, we investigate the distributional properties of n-words. N-words can express sentential negation (1-a). In combination with other n-words a single negation reading (SN) is possible (1-b). With a clause-mate negative marker (NM) *pas*, n-words trigger a DN reading (1-c).²

- (1) a. Jean n'a parlé à **personne**. [SN]
Jean NE has talked to nobody
'Jean hasn't talked to anyone.'
- b. **Personne** n'a **rien** dit. [SN,DN]
nobody NE has nothing said
'Nobody said anything.' [SN] or: 'Nobody said nothing.' [DN]
- c. **Personne** n'est pas venu. [DN]
nobody NE is not come
'Nobody did not come.'

¹The negative adverb *pas* occurs in comparatives as in (i). Wilmet (1997) uses this observation to argue that comparatives are, to a certain extent, negative.

- (i) a. Il faut avoir l'esprit plus libre que je ne l'ai **pas**.
it needs have the spirit more free than I NE it have NM
'One must have a freer spirit than I do.' (Racine, after Wilmet 1997, p. 513)
- b. Il est plus instruit que je suis **pas**.
he is more instructed than I am NM
'He is better instructed than I am.' (Sturm, 1981, p. 24)

²Non-European varieties of French show single negation readings for sentences such as (1-c). See Sections 5.2.1 and 6 for a more detailed discussion.

To summarize the basic facts of French, n-words display optional NC, but the negative marker *pas* does not participate in NC.

Negation in Polish represents a second typological class.³ The examples in (2) show that sentential negation in Polish typically requires a pre-verbal negative marker *nie* (Kupść, 2000, ta; Kupść and Przepiórkowski, 2002). N-words must co-occur with the negative marker, and only a single negation reading is possible (2-b).

- (2) a. Janek *(nie) pomaga **nikomu**. [SN]
 Janek (NM) helps nobody
 'Janek doesn't help anybody.'
 b. **Nikt** *(nie) pomaga ojcu. [SN]
 nobody (NM) helps father
 'Nobody helps his father.'

More than one n-word can occur in the same clause (3). In these cases, again, only a single negation reading is available.

- (3) a. **Nic** **nikomu** *(nie) powiedziałem. [SN]
 nothing.GEN nobody.DAT (NM) I-told
 'I didn't tell anybody anything.'
 b. **Nikt** **nigdy** **nikogo** **niczym** *(nie) uszczęśliwił. [SN]
 nobody never nobody.GEN with nothing (NM) made happy
 'Nobody has ever made anybody happy with anything.'

The obligatoriness of the NM in Polish makes it difficult to distinguish n-words from *negative polarity items* (NPI). NPIs are expressions that cannot occur in affirmative statements. Many languages have a group of indefinite NPIs, such as English *anything* or *a word*. Błaszczak (1999) argues for an analysis of Polish n-words as NPIs. Richter and Sailer (2004a) provide counter-arguments in favor of the inherent negativity of Polish n-words. In particular, in non-verbal projections, n-words can express negation, whereas uncontroversial corresponding NPIs cannot. The examples in (4) show that the NPI *słowo* (*a word*) is ungrammatical if there is no (potentially elided) licensing negation in short answers, whereas the n-word *żaden* (*no*) can occur without such licensing, (5). The genitive of negation in (4) may provide evidence for the presence of an elided verbal negation; accusative case excludes this as a straightforward option.

- (4) Powiedział coś? *Słowo./ Słowa./ Słowa nie powiedział.
 he said something Word.ACC/ Word.GEN/ Word.GEN NM he said
 'Did he say something/anything? Not even a word./ He did not say even a word.'
 (5) Ile przeczytałeś książek? **Żadną**./ **Żadnej**.
 How many you read books? None.ACC./ None.GEN.
 'How many books have you read? None'

³See Richter and Sailer (2004a,b) for an in-depth discussion of the Polish data and a review of the relevant literature.

In German the negation particle *nicht* and n-words (*niemand* (*no one*)) always express negation and never enter an NC relation. The data in (6) are syntactically parallel to those in (1).

- (6) a. Hans sprach mit **niemandem**. [SN]
 Hans talked with no one
 b. **Niemand** sprach mit **niemandem**. [DN]
 no one talked with no one
 c. **Niemand** kam nicht. [DN]
 no one came not
 ‘No one didn’t come.’

These data show that n-words in German do not enter NC. The negative marker *nicht* need not be present and does not enter into NC.

So far we have considered the negation systems of three types of languages. In all three languages, n-words must be considered as inherently negative. Nonetheless the interpretation of sentences with n-words and the possibility of their co-occurrence with other n-words and with the negative marker differ. On the basis of the interpretation of clauses with more than one n-word, we call French an *optional NC language*, Polish an *obligatory NC language*, and German a *no-NC language*. Giannakidou (2005) gives a typologically oriented overview over NC. According to her, optional NC is attested in Romance languages (Italian, Catalan), obligatory NC is found in the Slavic languages, but also in Greek, Hungarian, Rumanian and Japanese. English and Dutch are no-NC languages — at least in their standard variety.

2.2 Exceptions in French and German

In addition to the core data of the previous section, both French and German have a number of exceptional n-words.⁴ French *mot* expresses negation (7-a). In contrast to *personne* a DN reading is not possible in combination with n-words (7-b) and the combination with *pas* is ungrammatical (7-c).⁵ Surprisingly the German n-word *Dreck* behaves in exactly the same way in (8-a) and (8-b).⁶

- (7) a. Jean n’a dit **mot**. [SN]
 Jean NE has said word
 ‘Jean said nothing.’
 b. **Personne** n’a dit **mot**. [SN]
 nobody NE has said word
 ‘Nobody said anything.’

⁴We have not been able to find exceptional n-words in Polish so far.

⁵As pointed out to us by Olivier Bonami and Gilles Boyé, for many speakers (7-a) is not grammatical, whereas the other two sentences in (7) are. For these speakers *mot* is an NPI, similar to English *a word* or Polish *słowo* (*a word*) in (4).

⁶Some German speakers reject (8-c), which we find fully grammatical.

- c. *Il ne dit pas **mot**.
he NE says not word
- (8) a. Das geht dich **einen Dreck** an. [SN]
this concerns you a dirt PART
'This is none of your business.'
- b. Das geht **niemanden einen Dreck** an. [SN]
this concerns no one a dirt PART
'This is no one's business.'
- c. Das geht dich **keinen Dreck** an. [SN]
this concerns you no dirt PART
'This is none of your business.'

Notice that while *mot* and *Dreck* behave like n-words with respect to the truth conditions of the respective sentences, they are severely constrained with respect to the verbs they can combine with. In French the original lexical meaning of specialized n-words such as *mot* is an important factor: *mot* (literally: *word*) can only combine with verbs of saying.

- (9) a. Jean n'a dit **rien** du tout/ **mot**.
Jean NE has said nothing at all/ word
- b. Jean n'a acheté **rien** du tout/ ***mot**.
Jean NE has bought nothing at all/ word

Similarly, German *Dreck* only combines with a restricted number of verbs, verbs of intellectual concern such as *kümmern*, *scheren* (both meaning *care* or *concern*), or *interessieren* (*interest*).

- (10) a. Das schert/ interessiert mich **einen Dreck**/ gar nicht.
this concerns/ interests me a dirt/ not at all
'I don't care about this at all.'
- b. Das gefällt mir ***einen Dreck**/ gar nicht.
this pleases me a dirt/ not at all
'I don't like this at all.'

3 Precursors and Theoretical Prerequisites

3.1 NC with Truth Conditional Semantics in HPSG

There are a number of previous studies on NC in HPSG which link syntax to a truth conditional semantic analysis. De Swart and Sag (2002) provide an HPSG analysis of NC in terms of the lexical retrieval of quantifiers. Lexical retrieval is combined with the option of forming a polyadic quantifier, i.e. merging a sequence of expressions of the form $\neg\exists x_i$ into a single quantifier $\neg\exists x_1 \dots x_n$. A language-specific parameter will determine whether such an absorption is possible. DeSwart (2006) uses this syntactic framework to provide an optimality theoretic account of the characteristic interpretation strategies in a number of languages. This theory

captures the general patterns (NC/non-NC) of the languages, but it remains unclear how to incorporate lexical idiosyncrasies which contradict the general pattern of a language in this analysis.

Richter and Sailer (1999) discuss a set of data similar to those we investigate here. Their analysis, formulated in terms of a traditional Ty2 semantics using the lambda calculus and type shifting rules for the semantic combinatorics, focuses on the idiosyncrasies of the French data and models all of French negation in terms of a lexical ambiguity of n-words and idiosyncratic collocational restrictions for each reading of the n-words. While this approach describes both the general pattern and the idiosyncratic data, it fails to capture typological generalizations and a distinction between the general case and exceptions. This distinction is, however, clearly present in the data.

Richter and Sailer (2004a) present an analysis of Polish as a strict NC language. The analysis uses Lexical Resource Semantics (LRS) and exploits the possibility that two items may contribute the same negation to the logical form of a clause. They enforce strict NC by a language-specific principle saying that, in Polish, every verbal projection may have at most one negation in its logical form. This analysis accounts for one particular general pattern of NC in a fairly elegant way. However, it has not been shown how different NC patterns ranging from obligatory to impossible concord can be accommodated.

3.2 Lexical Resource Semantics

Following Richter and Sailer 2004a, our semantic interpretations will be couched in terms of LRS. LRS crucially allows us to use (1) a semantic combinatorics different from the lambda calculus, (2) techniques of scope underspecification, (3) identity constraints for (pieces of) semantic representations, and (4) expressions of Ty2 as logical representations.

In LRS the semantic information of a sign is encoded in its L(OGICAL-)F(ORM) value. The value of this attribute contains the following two attributes:⁷ PARTS lists all subexpressions that are contributed by a sign. The EX(TERNAL-)C(ONTENT) is the logical form of a phrase. The combinatoric principles determine that the PARTS list of a phrase is the concatenation of the daughters' PARTS lists. Furthermore, the EXC value of an utterance consists exactly of the expressions on the utterance's PARTS list.

3.3 A Collocation Module

Richter and Sailer (1999) use a collocation module to account for n-words in general. Soehn (2006) modifies this module in a theory of idiomatic expressions and integrates it with an LRS semantics. A sign has a list-valued attribute COLL. Collocationally restricted items have a non-empty COLL value, which may contain

⁷LRS uses some more attributes, which, however, do not play a role in this paper.

various *barrier* objects indicating the syntactic domain in which their context requirements must be satisfied. For our data, this will always be the smallest complete clause containing a given lexical item. Barrier objects have several attributes which are used to specify (local) syntactic or semantic properties that the relevant barrier must have, such as LOC-LIC for its *local* value and LF-LIC for properties of its logical form.

4 Analysis I: The Typological Patterns

4.1 NC Universals

It has been argued in the literature that the conceptually most attractive analysis of the data is one which assumes a single lexical entry for any given n-word and characterizes their occurrence restrictions in terms of entailment properties of the admissible contexts of the n-word (see for example Giannakidou (1997)). The data above permit a treatment with a single lexical entry for each n-word only under the assumption of negation agreement. Consider sentences with the n-word *mot* but without another n-word. Then the only potential overt source of a negation in the clause is the n-word *mot*. Negation must, thus, be part of its semantic contribution, which in turn must be licensed by the lexical entry of the word. When *mot* occurs together with *personne*, we would erroneously predict the absence of an SN reading unless we assume negation agreement. The same observation holds for the other n-words. Clearly negative instances like the examples (1), (5) and (6-a) above force us to assume that negation is part of the semantic contribution of n-words in all three languages under consideration. In (11) we state the common properties of all n-words considered in the present paper.

(11) Schematic lexical description of an n-word:⁸

$$\left[\begin{array}{l} \text{PHON } \langle \text{personne/nikt/niemand} \rangle \\ \text{SYNSEM NP} \\ \text{LF } \left[\begin{array}{l} \text{EXC } \boxed{1} \exists x(\alpha \wedge \beta) \\ \text{PARTS } \langle x, \boxed{1}, \text{human}'(x), \neg\gamma \rangle \end{array} \right] \end{array} \right] \text{ and human}'(x) \triangleleft (\text{is a component of}) \alpha$$

and $\boxed{1}$ is a component of γ

Given the characteristics of LRS mentioned in Section 3.2, optional negation agreement is available as a basic option of the semantic combinatorics: Each n-word contributes negation (\neg), but n-words can agree, i.e. they may contribute the same semantic negation to an utterance.

Our discussion of the core data has revealed that the negative marker also contributes negation in all three languages. In (12) the relevant semantic contribution of the NM is sketched. Note that syntactically the NMs differ considerably: German *nicht* is a simple adverb, Polish *nie* is a verbal prefix and French *pas* is an

⁸Greek letters in the descriptions refer to subterms which are not specified in more detail. Occasionally we write $\alpha \triangleleft \beta$ to indicate that the Ty2 expression described by α is a subterm (component) of the Ty2 expression described by β . PARTS lists are somewhat simplified throughout this paper.

adverb to VP or a complement of the verb (Kim and Sag, 2002).

(12) Schematic lexical description of a negative marker:⁹

$$\left[\begin{array}{l} \text{PHON } \langle \text{pas/nie/nicht} \rangle \\ \text{SYNSEM } [\text{LOC CAT HEAD MOD LOC CAT HEAD } \textit{verb}] \\ \text{LF } [\text{PARTS } \langle \neg \delta \rangle] \end{array} \right]$$

Based on these sketches of lexical entries we can now look at the three configurations discussed in Section 2.1: a sentence with only an n-word, a sentence with an n-word and an NM, and a sentence with two n-words. In all cases, we will show what the possible interpretations are if we do not impose any language-specific constraints. In Section 4.2 we will introduce the principles which will restrict the range of readings to those which are actually attested in each language.

In (13-a) a sentence with one n-word and no NM is given in the three languages under discussion. In (13-b) the semantic contributions of the words in the sentence, i.e. their PARTS lists are stated.

- (13) a. (i) Jean n’a parlé à **personne**. (French)
 J. talked with nobody
 (ii) *Janek rozmawiał z **nikim**. (Polish)
 (iii) Hans sprach mit **niemandem**. (German)
 b. Semantic contributions of the words:
 n-word: $[\text{PARTS } \langle x, \exists x(\alpha \wedge \beta), \text{human}'(x), \neg \gamma \rangle]$
 proper name: $[\text{PARTS } \langle j \rangle]$
 verb: $[\text{PARTS } \langle \text{talk}'(j, x) \rangle]$

Due to the combinatoric principles of LRS, the PARTS list of the sentences in (13) contains exactly the elements of all the PARTS lists of the words in the sentence. The resulting list is indicated in (14).

- (14) PARTS list of the sentences in (13):
 $[\text{PARTS } \langle x, \exists x(\alpha \wedge \beta), \text{human}'(x), \neg \gamma, j, \text{talk}'(j, x) \rangle]$

We know that the logical form of the sentences must be composed of exactly the expressions in (14). However, the list does not explicitly encode the relative embedding of these expressions. For example, we do not know from looking at (14) whether $\text{human}'(x)$ occurs in the restrictor or in the scope of the existential quantifier, i.e., whether it is a component of α or β . This information is partially specified in the lexical entries as well as in the combinatorial principles of LRS. In (15) the relevant restrictions are indicated, together with their source.

⁹Since Polish *nie* is a verbal prefix, (12) has to be re-interpreted in this case as describing the semantic contribution of *nie* to Polish negated verb forms. The MOD feature does not play a role in Polish, except for indicating that *nie* modifies the semantics of verbs in morphology.

- (15) Subterm constraints on the semantic contributions:
 $\text{human}'(x) \triangleleft \alpha$ (lexical entry of the n-word)
 $\exists x(\dots) \triangleleft \neg\gamma$ (lexical entry of the n-word)
 $\text{talk}'(j, x) \triangleleft \beta$ (syntactic combination of verb + n-word)

Only the logical form in (16) consists exactly of the expressions on the PARTS list and at the same time satisfies these constraints.

- (16) Potential reading: $\neg\exists x(\text{human}'(x) \wedge \text{talk}'(j, x))$

The second type of sentences that we want to discuss contains one n-word and an NM. The examples for our three languages are given in (17-a). In (17-b) the lexical contribution of the NM is given. The contribution of the other words is identical to that in (13-b) above.

- (17) a. (i) Jean n'a pas parlé à **personne**. (French)
 J. NE has NM talked with nobody
 (ii) Janek nie rozmawiał z **nikim**. (Polish)
 (iii) Hans sprach mit **niemandem** nicht. (German)
 b. Semantic contributions of the words:
 see (13-b)
 n-marker: $[\text{PARTS } \langle \neg\delta \rangle]$

Collecting these expressions, we arrive at the PARTS list in (18).

- (18) PARTS list of the sentences in (17):
 $[\text{PARTS } \langle x, \exists x(\alpha \wedge \beta), \text{human}'(x), \neg\gamma, \neg\delta, j, \text{talk}'(j, x) \rangle]$

In order to deduce the logical forms of the sentences, we have to consider the subterm constraints contributed by the lexical entries and imposed by their syntactic combination. The relevant restrictions are collected in (19).

- (19) Constraints on the semantic embedding:
 $\text{human}'(x) \triangleleft \alpha$ (lexical entry of the n-word)
 $\exists x(\dots) \triangleleft \neg\gamma$ (lexical entry of the n-word)
 $\text{talk}'(j, x) \triangleleft \delta$ (combination NM + verb)
 $\text{talk}'(j, x) \triangleleft \beta$ (syntactic combination verb + n-word)

In (20) we indicate the logical forms which are compatible with these conditions.

- (20) Potential readings:¹⁰
 a. $\neg\neg\exists x(\text{human}'(x) \wedge \text{talk}'(j, x))$
 b. $\neg\exists x(\text{human}'(x) \wedge \neg\text{talk}'(j, x))$ [DN]
 $(= \forall x(\text{human}'(x) \rightarrow \text{talk}'(j, x)))$

¹⁰Note that $\neg\exists x\neg(\text{human}'(x) \wedge \text{talk}'(j, x))$ is excluded due to background assumptions about representing quantifiers syntactically as generalized quantifiers, i.e. $\exists x(\alpha \wedge \beta)$ as $\exists(x, \alpha, \beta)$.

$$c. \quad \neg \exists x(\text{human}'(x) \wedge \text{talk}'(j, x)) \text{ [SN]}$$

Out of the three potential readings, (20-c) is the single negation reading which is attested for the French and the Polish sentence in (17). It may arise since nothing enforces that $\neg\gamma$ (contributed by the n-word) and $\neg\delta$ (contributed by the NM) be distinct expressions.

The logical form in (20-b) is the double negation reading that we reported for German and French in Section 2.1. The logical form contains two negations. It arises if $\neg\gamma$ and $\neg\delta$ are extensionally distinct logical forms ($\gamma = \exists x(\dots)$, $\delta = \text{talk}'(j, x)$). The reading in (20-a) has not yet been mentioned in our discussion. It is available in German and French. However, it requires special stress patterns and comes with restrictions on word order in German. As a denial form it is used to reject a previous claim that John had talked to nobody (i.e. that $\neg \exists x(\text{human}'(x) \wedge \text{talk}'(j, x))$ is true).¹¹

Finally, we consider the sentence type with two n-words and no NM (21-a). In (21-b) we state the semantic contribution of the second n-word, which is analogous to that of the first n-word in (13-b).

- (21) a. (i) **Personne** n'a parlé à **personne**. (French)
 Nobody NE has talked to nobody
 (ii) ***Nikt** rozmawiał z **nikim**. (Polish)
 (iii) **Niemand** sprach mit **niemandem**. (German)
 b. Semantic contributions of the words:
 2nd n-word: $\text{[PARTS } \langle y, \exists y(\alpha' \wedge \beta'), \text{human}'(y), \neg\gamma' \rangle]$

The lexical semantic contributions add up to the PARTS list in (22).

$$(22) \quad \text{PARTS list of the sentences in (21):}$$

$$\left[\text{PARTS } \left\langle x, \exists x(\alpha \wedge \beta), \text{human}'(x), \neg\gamma, y, \exists y(\alpha' \wedge \beta'), \right. \right. \\ \left. \left. \text{human}'(y), \neg\gamma', \neg\delta, y, \text{talk}'(y, x) \right\rangle \right]$$

In addition to the embedding constraints in (15) we also know that $\text{human}'(y)$ must be in the restrictor of the second existential quantifier, i.e., it must be a component of α' , and we know that the scope of the quantifier must contain $\text{talk}(x, y)$, i.e. $\text{talk}'(x, y)$ must be a component of β' .

If, in addition, we assume that the subject takes scope over the direct object, we can derive three possible readings, given in (23).

- (23) Potential readings (assuming subject > object)
- $\neg \neg \exists y(\text{human}'(y) \wedge \exists x(\text{human}'(x) \wedge \text{talk}'(y, x)))$
 - $\neg \exists y(\text{human}'(y) \wedge \neg \exists x(\text{human}'(x) \wedge \text{talk}'(y, x)))$ [DN]
 - $\neg \exists y(\text{human}'(y) \wedge \exists x(\text{human}'(x) \wedge \text{talk}'(y, x)))$ [SN]

The double negation reading in (23-b) is attested for the French and the German sentence in (21). The single negation reading, (23-c), is found for French.

¹¹We are grateful to Danièle Godard for discussion of the French data.

The reading in (23-a) seems to be absent in all three languages. The problem is that there is a negation intervening between the negation and the existential quantifier contributed by the same n-word: If the first “¬” is contributed by the subject, then the “¬” of the direct object intervenes between the subject’s negation and its quantifier ($\exists x \dots$). If the first “¬” stems from the direct object, then the subject’s “¬” intervenes between the object’s negation and its quantifier ($\exists y \dots$). We can exclude this kind of linear intervention by adding the following line to the lexical specification of n-words in (11).¹²

(24) Intervention condition, to be added to the specification in (11):

$$\text{and not } \mathbf{E} \in \left(\neg \epsilon \triangleleft \gamma \text{ and } \exists x(\alpha \wedge \beta) \triangleleft \epsilon \right)$$

The condition in (24) says that there may not be an expression ϵ such that $\neg \epsilon$ is in the scope of the negation contributed by the n-word ($\neg \gamma$), and at the same time, the existential quantifier contributed by the n-word ($\exists x \dots$) is in ϵ .

In this section we showed that we can derive all and only the attested readings if we assume the lexical specifications for the n-words in (11) (augmented with (24)) and for the negative marker in (12), and apply the combinatorial principles of LRS without any further restrictions. We saw that not all of the resulting readings are available in all languages. In the next section we will present language-specific principles that will allow us to impose the correct restrictions for each language.

4.2 Typological Constraints

In this section we will present the general principles which determine the typological type of the negation system of each language. We will first look at Polish, then at German, and finally at French.

4.2.1 Polish

To enforce obligatory negative concord for Polish, Richter and Sailer (2004a) proposes the NEGATION COMPLEXITY CONSTRAINT, given in (25). Remember that in LRS, the EXCONT value is the logical form of a sign, the MAIN value is the main semantic constant contributed by the sign’s lexical head.

(25) The NEGATION COMPLEXITY CONSTRAINT:
For each sign, there may be at most one negation that is a component of the EXCONT value and has the MAIN value as its component.

With this principle we can rule out the double negation reading and the denial reading given in (20) for the Polish sentence in (17). In this sentence, the MAIN

¹²To distinguish clearly between the HPSG description language and the semantic representation language we use **not**, **and** and **E** for negation, conjunction and existential quantification in the former.

value is the constant **talk'**. In the hypothetical EXCONT values in (20-a) and (20-b) this constant is in the scope of two negations.

To guarantee the presence of the NM in negated verbal projections, we must invoke a second language-specific principle for Polish. We called this principle the NEG CRITERION, due to its similarity in effect to the Neg Criterion of Haegeman and Zanuttini (1996).

(26) The NEG CRITERION:

For every verb, if there is a negation in the EXCONT value of the verb that has scope over the verb's MAIN value, then that negation must be an element of the verb's PARTS list.

Since the Polish NM *nie* is a verbal prefix, its semantic contribution is part of the semantic contribution of a negated verb. If a sentence contains an n-word and a negated verb, as the Polish sentence in (17), the NEG CRITERION is met. In (13) and in (21), however, there is no negation in the semantic contribution of the verb. Despite its absence in the verb, in all the potential logical forms of the complete sentence the verb's MAIN value (**talk'**) is in the scope of a negation. Thus, the NEG CRITERION correctly excludes these sentences. Note that if we add a pre-verbal negation in sentence (21), the sentence becomes grammatical. Due to the NEGATION COMPLEXITY CONSTRAINT it has only a single negation reading.

4.2.2 German

For Polish the constraint in (25) prevents double negation readings. For German, a non-NC language, we need a constraint that prohibits negation agreement. This constraint is given in (27).

(27) NEGATION FAITHFULNESS CONSTRAINT (German, Dutch, English):

a. In every phrase: there is no element of the form $\neg\alpha$ which is on the PARTS list of both the head-daughter and the nonhead-daughter.

$$b. \quad phrase \Rightarrow \left(\begin{array}{c} \left[\begin{array}{l} \text{H-DTR LF PARTS } \boxed{A} \\ \text{N-DTR LF PARTS } \boxed{B} \end{array} \right] \\ \text{and not } E_{\boxed{A}} E_{\alpha} \left(\begin{array}{c} \boxed{1} = \neg\alpha \\ \text{and member}(\boxed{1}, \boxed{A}) \\ \text{and member}(\boxed{1}, \boxed{B}) \end{array} \right) \end{array} \right)$$

Let us reconsider the sentences with two words which contribute a negation, i.e. sentences (17) and (21). Their single negation readings in (20-c) and (23-c) arise by the identification of the negation contributed within the VP (either by the negative marker or by the n-word in object position) and the negation contributed by the subject. As an immediate consequence, when the subject combines with the VP, both constituents have a negation in their PARTS list. To derive a single negation reading, these negations must be identical, i.e. there must be an expression $\neg\alpha$

which is on the head-daughter's and on the nonhead-daughter's PARTS list. This is, however, what the constraint in (27) excludes.

4.2.3 French

For French n-words of the *personne*-type nothing needs to be said, as LRS allows for negation agreement but does not enforce it. This means that the ambiguity of a sentence with two n-words such as (21) is correctly predicted. At the moment we cannot yet exclude the single negation reading of a sentence with an n-word and the NM (i.e. of the type in (17)). This has to be postponed until the next section, where we will derive it from lexical properties of *pas* in European French.

Olivier Bonami and Gilles Boyé (p.c.) brought to our attention a number of interesting n-words in French slang, such as *que dalle*, *que tchi* and *oualou* (all meaning *nothing*), which are beyond the well-known core of French n-words.¹³ According to a preliminary google search and introspective judgements of a small number of native speakers, these n-words pattern exactly like other n-words in French: They license negative polarity items (28-b), they show negative concord with other n-words (29), and they express a double negation when combined with *pas* (30). We did not find an instance of a double negation reading with *quedalle* and another n-word, though.

- (28) a. *Je fous toute sorte de chose pendant les vacances.
I made all sorts of things during the holidays
b. On foutait rien/ que dalle/ que tchi/ oualou
One made nothing
'We did nothing.'

- (29) (internet data)
a. mais si on va dans ce sens là, plus **personne** fait
but if one goes in this direction there no more nobody does
quedalle
nothing
'but if one goes in this direction, nobody does anything anymore.'
(found by O. Bonami)
b. en réalité ces initiatives n'apportent absolument **que dalle** à
in reality these initiatives NE bring absolutely nothing to
personne ...
nobody
'In reality, these initiatives don't serve anything at all to anybody.'

- (30) C'est pas quedalle
that's NM nothing
'That's not nothing.'

¹³Their non-standard status is also documented by (i) their high frequency of occurrence without *ne* and (ii) the considerable amount of orthographic variants (which include for *que dalle*: *quedalle*, *kedal*, *que le dail*, for *que tchi*: *ketchi*, *keutchi*, for *oualou*: *waloo*, *walloo*, *walou*).

While these n-words deserve a more systematic investigation, our preliminary survey suggests that they follow the pattern of the well-studied n-words in French.

4.3 Summary

In this section we demonstrated that we can assume the same semantic specifications for n-words and negative markers in French, Polish, and German. The typological differences in the negation systems are derived from language-specific restrictions on the mutual compatibility of negative items in a sentence.

In our theory optional NC is the simplest case, which is typologically correct. Strict NC can be enforced and might even be preferred because it leads to less complex logical forms. For the rare cases of non-NC languages a principle like (27) can account for the general pattern. Thus, these languages have more complex grammars than NC languages, which may explain their typological markedness.

5 Analysis II: Exceptional N-Words

While the negation agreement behavior of *personne/nikt/niemand*-type n-words follows from the architecture of LRS and general typological principles, our account of the n-words *mot* and *Dreck* in (7) and (8), and of the negative adverb *pas* is lexicalized and treats them as idiosyncratic items. Their lexical entries contain collocational restrictions which exclude some of the readings we expect to find according to the general principles.

5.1 Collocationally Restricted N-Words

In Section 2.2 we showed that French *mot* (*word*) and German *einen Dreck* (*a dirt*) are inherently negative, exhibit obligatory NC with other negative items in the sentence, and are restricted to co-occur only with a small number of verbs. French has a number of n-words similar to *mot*: The n-word *goutte* (*drop*) co-occurs with verbs of drinking, but also with verbs of perception (*voir*, *entendre* (*see*, *hear*)), or comprehension (*comprendre*, *connaître* (*understand*, *know*)). The n-word *mie* (*crumb*), the most archaic of the three, is attested with *écouter* (*listen*) and *attendre* (*wait/expect*). Our brief overview shows that the verbs with which each of these n-words co-occurs are not fully predictable from the literal meaning of the n-words.

It is worth noting that the negation marker *pas* was historically just one more of these specialized n-words. Motivated by its literal meaning (*step*), it used to combine preferably with verbs of movement. In Section 4.2.3 we observed that the interaction of *pas* with n-words does not follow completely from the general principles of the negation system of European French. For other varieties such as Canadian French (Acadian), reported in Richter and Sailer (1999), *pas* is fully integrated in the optional concord system. There we find both single and double negation readings for sentences which contain an n-word and *pas* ((1-c), (17)). We

conclude that in European French, *pas* has a regular grammatical meaning but, nonetheless, is (still) not free from idiosyncratic collocational restrictions.

In Section 2.2 we showed that the exceptional German n-word *einen Dreck* is similar to French *mot*: It is inherently negative, it does not lead to double negation readings and it is collocationally restricted to a small class of verbs. German has a number of other such specialized n-words (Sailer *ta*), including *den/ einen Teufel* (*the/ a devil*) or *einen feuchten Kehricht* (*a damp dust*), which collocate with roughly the same class of verbs attested for *einen Dreck*.

The phrase *einen Dreck* violates the general typological pattern of German, which excludes NC. Interestingly, we also find French n-words that go against the otherwise stable ban on single negation readings with *pas*. There is a third group of n-words in French, also mentioned in Richter and Sailer 1999, which includes *âme que vive* (*soul that lives*). This n-word behaves analogously to *mot*, but it can form a single negation reading with *pas*, excluding a double negation reading.

- (31) a. Il n'y a (pas) **âme qui vive** dans cet endroit désert. [SN]
 It NE there has NM a living soul in this place deserted
 'There isn't a living soul in this deserted place'
- b. **Personne** n'a **jamais** rencontré **âme qui vive** dans cet endroit
 nobody NE has never met a living soul in this place
 désert. [SN]
 deserted
 'Nobody has ever met anyone in this deserted place.'

5.2 Analysis

We are now ready to show that the behavior of the exceptional words can be captured using the collocation theory of Soehn (2006) outlined in Section 3.3. The necessary collocational restrictions will directly express the distributional idiosyncrasy we find: For *pas* it will refer to abstract items in the logical form, for *mot* it will mention the verb class. The lexical entries of *einen Dreck* and *âme qui vive* will be the most elaborate, reflecting their unusual behavior relative to the negation systems of French and German.

5.2.1 The Lexical Entry of *pas*

Since *pas* is a negative marker its semantic contribution is as described for negative markers in general in (12). However, we have to add a collocational restriction within its COLL value in which we specify that no other item may agree with it within the same clause. This enforces the DN reading in (1-c) and in (17).

- (32) Sketch of the relevant part of the lexical entry of *pas*:

$$\left[\begin{array}{l} \text{PHON } \langle \textit{pas} \rangle \\ \text{SYNSEM ADV} \\ \text{LF } [\text{PARTS } \langle \neg\delta \rangle] \\ \text{COLL } \left\langle \left[\begin{array}{l} \textit{complete-clause} \\ \text{LF-LIC } [\text{PARTS } \boxed{A}] \end{array} \right] \right\rangle \end{array} \right] \text{ and } \neg\delta \text{ occurs exactly once in } \boxed{A}.$$

5.2.2 The Lexical Entry of *mot*-type N-Words

The lexical entry of *mot* in (33) is consistent with (11). In addition, it contains a non-empty COLL value expressing that: (i) its collocational restrictions must be satisfied in the smallest complete clause containing *mot*; (ii) in this clause, *mot* must combine with a verb of saying (we use the attribute LISTEME from Soehn 2006 to express this); (iii) while *mot* contributes a negation, this negation may not be distinct from other negations in the same clause. Under this analysis, the incompatibility of *pas* and *mot* in (7-c) is an immediate consequence of the contradictory collocational requirements of the two items.

- (33) Sketch of the lexical entry of the exceptional n-word *mot*:

$$\left[\begin{array}{l} \text{PHON } \langle \textit{mot} \rangle \\ \text{SYNSEM NP} \\ \text{LF } \left[\begin{array}{l} \text{EXC } \boxed{1} \exists x(\alpha \wedge \beta) \\ \text{PARTS } \langle x, \boxed{1}, \textit{thing}'(x), \neg\gamma \rangle \end{array} \right] \\ \text{COLL } \left\langle \left[\begin{array}{l} \textit{complete-clause} \\ \text{LOC-LIC } [\text{CAT HEAD LISTEME } \textit{saying}] \\ \text{LF-LIC } [\text{PARTS } \boxed{A}] \end{array} \right] \right\rangle \end{array} \right] \\ \text{and } \textit{thing}'(x) \text{ is a component of } \alpha \text{ and } \boxed{1} \text{ is a component of } \gamma \\ \text{and if there is an element in } \boxed{A} \text{ of the form } \neg\delta, \text{ then } \delta = \gamma$$

5.2.3 The Lexical Entry of *einen Dreck*-type N-Words

At the surface the pattern of German *Dreck* in (8) is analogous to that of *mot*. However, we have to take into account that the negation systems of French and German are fundamentally different. French has optional NC, in German NC is impossible. We assume that *Dreck* is lexically specified as optionally introducing a negation. Collocationally it is just like *mot*. A clause-mate negation may not be distinct from the negation contributed by *Dreck*. This leads to the effect that *Dreck* does not contribute a negation in the context of a negative marker or an n-word.

This analysis also makes the right predictions for (34), in which there are two words, an n-word and a negative determiner, which contribute negation.

- (34) Das schert **niemanden keinen Dreck**. [DN]
 this concerns no one no dirt
 'No one does not care about this.'

In this case, the typological pattern of German is responsible for the double negation reading, and *Dreck* does not contribute a negative component.

(35) Sketch of the lexical entry of the exceptional n-word *Dreck*:

$$\left[\begin{array}{l} \text{PHON } \langle \textit{Dreck} \rangle \\ \text{SYNSEM NP} \\ \text{LF } \left[\begin{array}{l} \text{EXC } \boxed{\perp} \exists x(\alpha \wedge \beta) \\ \text{PARTS } \langle x, \boxed{\perp}, \textit{thing}'(x) \rangle (\oplus \langle \neg\gamma \rangle) \end{array} \right] \\ \text{COLL } \left\langle \begin{array}{l} \text{complete-clause} \\ \text{LOC-LIC } \left[\begin{array}{l} \text{CAT HEAD LISTEME } \textit{intell-concern} \\ \text{LF-LIC } \left[\text{PARTS } \boxed{A} \end{array} \right] \end{array} \right] \right\rangle \end{array} \right]$$

and $\textit{thing}'(x)$ is a component of α
and $\boxed{\perp}$ is a component of an expression $\neg\delta$ in \boxed{A}
and if there is an element in \boxed{A} of the form $\neg\delta$, then $\delta = \gamma$

To model n-words such as *âme qui vive* (a living soul) in (31), we assume a lexical entry which is like the one for *einen Dreck* in (35) with different local collocation requirements (i.e. a different LOC-LIC specification), but with identical logical form collocations: The PARTS list contains an optional negation operator (“ $(\oplus \neg\gamma)$ ”), and the logical form of the smallest clause containing the n-word must have a negation ($\neg\delta$), but this negation may not be distinct from $\neg\gamma$. It follows that whereas the version of *einen Dreck* without negation is required in German whenever another negative item occurs in the same clause, *âme que vive* only needs to be non-negative if it co-occurs with *pas*.

6 Summary

The theory of NC which we have developed in this paper has three layers. The universal core system is determined by the semantic combinatorics of LRS and the structure of the collocation theory. At the same time common lexical semantic specifications of the important words of the sentential negation system (n-words and negative markers) have been identified. Without additional assumptions, the core system delineates the same potential readings for French, Polish and German. In the second layer, the typological principles distinguish among the three language-specific typological classes of NC we saw. In the third layer, language internal idiosyncrasies, i.e. exceptions to the general typological class, are handled by exceptional collocation requirements of small classes of lexical items. As a result, we distinguish clearly between (i) the overall type of the language and (ii) lexical items with principle-governed versus idiosyncratic behavior. Previous approaches have not been able to combine these two aspects.

Our analysis distinguishes three typological classes of NC. French is the simplest case, since the core of n-words exhibits an unmarked behavior. Double negation readings and single negation readings with two n-words are optional. The system is unstable due to a collocationally restricted function word, the negative

marker *pas*. Polish is a pure NC language. Two language-specific principles ((25) and (26)) enforce obligatory NC and the presence of the verbal negative marker, *nie*, in negative sentences. The obligatory presence of the NM in Polish makes it non-trivial to distinguish n-words from NPIs in this system. German marks the other end of the scale. German forbids NC with a third language-specific principle, which we called NEGATION FAITHFULNESS CONSTRAINT (27).

The analysis makes the following predictions: (1) We expect that functional items in a language show fewer or no collocational properties. This is attested for French *pas*. Historically, *pas* was collocationally as restricted as *mot*. Its collocational restrictions today are more general than those of the *mot* type. However, the incompatibility of *mot* and *pas* still follows from their respective collocational requirements alone. In non-European variants of French, *pas* does not have a collocational restriction, i.e. it behaves according to the general principle and permits optional NC. (2) For non-NC languages, items which enforce “NC-like” interpretations have complex semantic contributions and collocational requirements. Thus, they are highly marked. In fact, we only find very few of them in German. They have not been noticed in the literature, and we conjecture that their overall frequency in languages is very low.¹⁴

At the heart of our analysis is the technique of enforcing, forbidding or permitting structural identity between (components of) signs in complex structures. In HPSG this is *the* single most important device of linguistic description. It is used to model agreement in the nominal domain, coindexation in Binding Theory and subject-verb agreement in the sentential domain. In LRS analyses of semantic phenomena, structural identity of semantic representations has been used before to model tense agreement in Afrikaans (Sailer 2004) and interrogative agreement in multiple wh-questions in German (Richter and Sailer 2001). In the present contribution we argued that a typologically oriented analysis of NC can exploit negation agreement to account simultaneously for (1) the dominance of NC or multiple negation in a given language and (2) the occurrence of lexically marked exceptions to each pattern. We integrated lexical exceptions in such a way that they are distinguished as special cases which need to be learned individually.

¹⁴Postal (2004) presents intriguing data on idiosyncratic English slang n-words/minimizers such as *squat*. They seem to mean *nothing* in isolation (i-a), and they don’t allow for a DN reading if a negative marker or an n-word is present ((i-b), (i-c)). However, in contrast to French *mot* and German *Dreck*, *squat* does not license NPIs (such as *in years* in (i-d)). Postal, thus, treats *squat* as being ambiguous between a zero-quantifier in (i-a) and an NPI in (i-b) and (i-c). We do not attempt to analyze *squat* in this paper, but the data supports the point that NC-like items in non-NC languages show clearly marked behavior.

- (i) a. Claudia saw squat. (= Claudia saw nothing.)
- b. Claudia didn’t see squat. (= Claudia didn’t see anything.)
- c. Nobody knows squat about your topic. (= Nobody knows anything . . .) (internet data)
- d. Helga has said nothing/ *squat to me in years.

References

- Błaszczak, Joanna (1999). *Investigations into the Interaction between the Indefinites and Negation*. Ph. D. thesis, Humboldt Universität Berlin. Published in: Akademie Verlag, Berlin, *Studia grammatica* 51, 2001.
- Cristea, Theodora (1971). *La structure de la phrase négative en français contemporain*. Société Roumaine de Linguistique Romane, Bucarest.
- Gaätone, David (1971). *Étude descriptive du système de la négation en français contemporain*. Droz, Genève.
- Giannakidou, Anastasia (1997). *The Landscape of Polarity Items*. Ph. D. thesis, Rijksuniversiteit Groningen.
- Giannakidou, Anastasia (2005). N-Words and Negative Concord. In M. Everaert, H. van Riemsdijk, R. Goedemans, and B. Hollebrandse (Eds.), *The Linguistics Companion*, Volume 3, pp. 327–391. Oxford: Blackwell Publishing.
- Grevisse, Maurice and Goosse, André (1993). *Le bon usage. Grammaire française* (13 ed.). Duculot.
- Haegeman, Liliane and Zanuttini, Raffaella (1996). Negative Concord in West Flemish. In A. Belletti and L. Rizzi (Eds.), *Parameters and Functional Heads: Essays in Comparative Syntax*, pp. 117–179. Oxford University Press.
- Kim, Jong-Bok (1996). *The Grammar of Negation: A Lexicalist Constraint-Based Perspective*. Ph. D. thesis, Stanford University.
- Kim, Jong-Bok and Sag, Ivan A. (2002). Negation without Head-Movement. *Natural Language and Linguistic Theory* 20, 339–412.
- Kupść, Anna (2000). *An HPSG Grammar of Polish Clitics*. Ph. D. thesis, Polish Academy of Sciences and Université Paris 7.
- Kupść, Anna (t.a). The Morphosyntax of Polish Verbal Negation: Towards an HPSG Account. In S. Brown and A. Przepiórkowski (Eds.), *Negation in Slavic*. Slavica Publishers.
- Kupść, Anna and Przepiórkowski, Adam (2002). Morphological Aspects of Verbal Negation in Polish. In P. Kosta and J. Frasek (Eds.), *Current Approaches to Formal Slavic Linguistics*, pp. 337–346. Peter Lang, Frankfurt. Proceedings of the Second European Conference on Formal Description of Slavic Languages, Potsdam, 1997.
- Muller, Claude (1991). *La Négation en Français. Syntaxe, sémantique et éléments de comparaison avec les autres langues romanes*. Librairie Droz, Genève.
- Postal, Paul (2004). *Skeptical Linguistic Essays*. Oxford: Oxford University Press.

- Richter, Frank and Sailer, Manfred (1999). A Lexicalist Collocation Analysis of Sentential Negation and Negative Concord in French. In V. Kordoni (Ed.), *Tübingen Studies in Head-Driven Phrase Structure Grammar*, Arbeitspapiere des SFB 340, Nr. 132, Volume 1, pp. 231–300. Universität Stuttgart and Universität Tübingen.
- Richter, Frank and Sailer, Manfred (2001). On the Left Periphery of German Finite Sentences. In W. D. Meurers and T. Kiss (Eds.), *Constraint-Based Approaches to Germanic Syntax*, pp. 257–300. Stanford: CSLI Publications.
- Richter, Frank and Sailer, Manfred (2004a). Basic Concepts of Lexical Resource Semantics. In A. Beckmann and N. Preining (Eds.), *ESSLLI 2003. Course Material I*, Volume 5 of *Collegium Logicum*, pp. 87–143. Kurt Gödel Society Wien.
- Richter, Frank and Sailer, Manfred (2004b). Polish Negation and Lexical Resource Semantics. In L. S. Moss and R. T. Oehrle (Eds.), *Electronic Notes in Theoretical Computer Science*, Volume 53, pp. 309–321. Elsevier.
- Rowlett, Paul (1998). A non-overt negative operator in French. *PROBUS International Journal of Latin and Romance Linguistics* 10(2), 185–206.
- Sailer, Manfred (2004). Past Tense Marking in Afrikaans. In C. Meier and M. Weisgerber (Eds.), *Proceedings of Sinn und Bedeutung* 8, Konstanz, pp. 233–248. www.ub.uni-konstanz.de/kops/volltexte/2004/1383.
- Sailer, Manfred (t.a.). Don't Believe in Underspecified Semantics. In O. Bonami and P. Cabredo Hofherr (Eds.), *Empirical Issues in Formal Syntax and Semantics*, Volume 6.
- Soehn, Jan-Philipp (2006). *Über Barendienste und erstaunte Bauklötze. Idiome ohne freie Lesart in der HPSG*. Frankfurt am Main: Peter Lang. Ph.D. thesis, Friedrich-Schiller-Universität Jena.
- Sturm, Joachim (1981). *Morpho-syntaktische Untersuchungen zur phrase négative im gesprochenen Französisch. Die Negation mit und ohne NE*. Verlag Peter Lang.
- de Swart, Henriëtte (2006). Marking and Interpretation of Negation: A Bidirectional Optimality Theory Approach. In R. Zanuttini, H. Campos, E. Herburger, and P. Portner (Eds.), *Cross-linguistic Research in Syntax and Semantics. Negation, Tense and Clausal Architecture*, pp. 199–218. Georgetown: Georgetown University Press.
- de Swart, Henriëtte and Sag, Ivan A. (2002). Negation and Negative Concord in French. *Linguistics and Philosophy* 25, 373–417.
- Wilmet, Marc (1997). *Grammaire critique du français*. Duculot.